AMENDMENTS TO THE TITLE

Please replace the title with the following amended title:

Method~~, Aparatus and Article For Data Reduction~~ For Determining the Likelihood of a Match Between Source Data and Reference Data

AMENDMENTS TO THE SPECIFICATION

After the paragraph beginning at page 12, line 25, which starts with "Fig. 4 is a flow chart", please add the following new paragraph:

Fig. 5 is a flow chart of an exemplary method of utilizing processed data in accordance with various embodiments of the present invention.

Please replace the paragraph beginning at page 1, line 12, which starts with "There are many instances" with the following amended paragraph:

There are many instances ~~is~~ in which a great deal of data is produced, and it is a lengthy and costly process to analyze the data. By way of example only, and not by way of limitation, in the broadcast industry, it is desirable to monitor the programs broadcast by stations to determine whether specific content, such as for example music or songs, have been broadcast.

Please replace the paragraph beginning at page 2, line 28, which starts with "Electronic monitoring methodologies offer" with the following amended paragraph:

Electronic monitoring methodologies offer advantages over manual systems such as lower operating costs and reliability. One type of electronic monitoring methodology requires insertion of specific codes into broadcast information before the information is transmitted. The electronic monitoring system can then recognize a song, for ~~exampl~~ example, by matching the received code with a code in a reference library. Such systems suffer from both technical and legal difficulties. For example, such a coding technique requires circuitry, which is expensive to design and assemble and which must be placed at each transmitting and receiving station. Legal

difficulties stem from the adverse position of government regulatory agencies toward the alteration of broadcast signals without widespread acceptance thereof by those in the broadcast industry.

Please replace the paragraph beginning at page 3, line 23, which starts with "The apparatus and method described in the above" with the following amended paragraph:

The apparatus and method described in the above mentioned United States Patent No. 5,437,050 has met with commercial success and has in large part met the needs of many segments of the broadcast industry. This approach is based on the discovery that the broadcast information on which recognition is based lies in the narrow frequency bands associated with the semitones of the music scale, rather than in the continuum of audio frequencies or in other sets of discrete frequency bands. It is also based on the principle that the set of semitones that have energies above a threshold amount at each ~~instant~~ instance provide sufficient information for recognition, and that it is not necessary to use the absolute energies of all frequencies for recognition.

Please replace the paragraph beginning at page 4, line 6, which starts with "Thus, United States Patent No. 5,437,050" with the following amended paragraph:

Thus, United States Patent No. 5,437,050 provides an apparatus and a method of recognizing broadcast information, including the steps of receiving broadcast information, the broadcast information being in analog form and varying with time; converting the broadcast information into a frequency representation of the broadcast information; dividing the frequency representation into a plurality of separate frequency bands (generally 48 bands over four octaves);

determining a magnitude of each separate frequency band of the digital sample; and storing the magnitudes. The method of recognizing broadcast information also includes the steps of performing a significance determination a plurality of times, the significance determination including the steps of generating a magnitude of each separate frequency band, using a predetermined number of previously stored magnitudes for each respective frequency band; storing the magnitudes; and determining a significance value, using a predetermined number of previously stored magnitudes for each respective frequency band. The method of recognizing broadcast information further includes the steps of comparing the significance value to the most recently generated magnitude of each separate frequency bands generating a data array, the data array having a number of elements equal to the number of separate frequency bands, the values of the elements being either binary 1 or binary 0 depending on the results of the comparison; reading a reference data array, the reference data array having been generated from reference information; comparing the data array to the reference data array; and determining, based on the comparison, whether the broadcast information is the same as the reference information.

Please replace the paragraph beginning at page 7, line 22, which starts with "The present invention may be thought of as a lossy data" with the following amended paragraph:

The present invention may be thought of as a lossy data reduction technique. A series of input frames is each comprised of a set of N scalar values which may represent, for example, amplitude, magnitude or intensity of some characteristic of an original signal S. The nature of signal S, the choice of N characteristics, and the decomposition of signal S into input frames may take many different forms, and may be considered as independent of the invention. The input frames have a sampling rate Sr. The invention produces a series of output frames with a

sampling rate of that of the input frames divided by W ~~(that is Sr/W))~~ (that is Sr/W), where W is an averaging window, represented by a whole number greater than one, and wherein each output frame comprises N bits.

Please replace the paragraph beginning at page 8, line 28, which starts with "The resulting output has the following" with the following amended paragraph:

The resulting output has the following useful characteristics. First, it is a much smaller amount of data than the input. Specifically, the resulting size of the ~~data~~ output data is 1/(Nbits*W) times the size of the input data, where Nbits is the number of bits used to represent each of the scalar N values. Typical values of Nbits and W are 32 and 5 respectively, yielding a 1:160 reduction.

Please replace the paragraph beginning at page 9, line 6, which starts with "In addition, the bits set" with the following amended paragraph:

In addition, the bits set to 1 in the output data tend to represent the most salient characteristics of the input signal during the time period covered by each output frame. Further, a constant number of bits, Y, is set in each frame, making the resulting data and its properties amenable to straightforward analysis. The resulting output is more robust, in the sense that it is less impacted by noise, transients, and distortions than a conventional averaging technique. The output data can be used in many of the applications that would require the input data. These applications include signal comparison, feature detection, pattern recognition, anomaly detection, trend analysis, and etc. A significant increase in processing speed is provided, due to the

reduction in the amount of data that must be processed, and the fact that bit comparison operations can be used to process the data.

Please replace the paragraph beginning at page 10, line 9, which starts with "In the broadcast recognition" with the following amended paragraph:

In the broadcast recognition application, the method and apparatus divide the signal into a series of frames; for each frame, divide a spectrum of the signal into a series of frequency segments; determine which of a number of frequency segments of the series of frequency segments have the largest amplitudes; set a value of zero for all of the frequency segments other than the number having the largest amplitudes; set a value representative of amplitude for the frequency segments having the largest amplitudes, average respective values, for a series of frames, to produce a series of average values; select a number of the average values which are largest average values; and produce the digital representation by setting bits to a first binary value for the selected number of the average values, and to a second binary value for all other average values. The averaging of respective values, for a series of frames, to produce a series of average values includes averaging the values of zero.

Please replace the paragraph beginning at page 10, line 28, which starts with "The number of frequency segments" with the following amended paragraph:

The number of frequency segments of the series of frequency segments having largest amplitudes in the spectrum of the frame may be a predetermined, fixed number. The number of the average values having the largest average values that are selected may also be a

predetermined, fixed number. Preferably the first binary value is one and the second binary value is zero.

Please replace the paragraph beginning at page 11, line 4, which starts with "Determining which of a number" with the following amended paragraph:

Determining which of a number of frequency segments of the series of frequency segments have the largest amplitudes in the spectrum of the frame comprises performing a Fourier transform on the signal. Preferably, a Discrete Fourier Transform is used.

Please replace the paragraph beginning at page 11, line 21, which starts with "The invention is also directed" with the following amended paragraph:

The invention is also directed to a method for determining the likelihood of a match between a first set of data having Y of N bits set equal to a first binary value and a remainder of the bits set equal to a second binary value, and a second set of data also having Y of N bits set equal to a first binary value and a remainder of the bits set equal to a second binary value. This method comprises determining the general probabilities of Y of N bits in the first set of data and in the second set of data being the same; and heuristically processing the probabilities to produce a series of match values based on the number of respective bits in the first set of data and in the second set of data that are identical. The heuristic processing may comprise assigning a match of n out of Y values a value of 1; normalizing remaining values to the value of 1 to produce resulting numbers; multiplying the resulting numbers by a constant to produce multiplied numbers; and subtracting the multiplied numbers from 1 to produce the match values. It may

further comprise setting match values greater than a predetermined value to values substantially equal to 1.

Please replace the paragraph beginning at page 15, line 7, which starts with "Referring to Fig. 2C" with the following amended paragraph:

Referring to Fig. 2C, in a sixth column, the average value for each row is calculated, with the values set to zero averaged in as such. Then the Y largest average ~~value~~ values are selected. In this example Y=3, so that the largest three values are shown in bold-faced type. An Output is produced (column at the right) wherein the bit value for the selected Y largest values is set equal to binary 1. The remainder of the bit values are set equal to binary 0. The data reduction has been completed, with Y of the original values set to binary 1. Thus, as discussed in more detail below, the number of bits having a binary 1 value in the output of the data reduction process is constant, regardless of the precise nature of the data provided by the data source 10 (Fig. 1).

Please replace the paragraph beginning at page 20, line 3, which starts with "The probability of having" with the following amended paragraph:

Fig. 5 illustrates determining the general probabilities of Y of N bits being the same in various embodiments of the present invention, as indicated at step 100. The probability of having n bits match in any fr $\frac{\binom{8}{n}\binom{40}{8-n}}{\binom{48}{8}}$

Please replace the paragraph beginning at page 20, line 16, which starts with "Match values are assigned" with the following amended paragraph:

Fig. 5 illustrates heuristically processing the general probabilities in an embodiment of the present invention, as indicated at steps 102, 104, 106, 108, and 110. Match values are assigned heuristically with deference to and consideration of the probabilities of random data having that many bits match. As indicated in step 102, n out of Y values may be selected. The raw probabilities, calculated as in ~~equation~~ the equation above, for n out of eight bits matching are {.2038, .39, .284, .097, .017, .0014, .000058, .00000085, .0000000027}. The sum of these probabilities, if taken to the limit of accuracy, totals 1. A match of six out of eight may be taken as a baseline of a probability of 1.0 as indicated at 104. The values then become:

1 - 6,500
2 - 4,733
3 - 1,616
4 - 283
5 - 23
6 - 1
7, 8 - better than 1

Please replace the paragraph beginning at page 21, line 8, which starts with "The square roots of the" with the following amended paragraph:

Referring still to Fig. 5, the ~~The~~ square roots of the resulting numbers are taken, they are multiplied by 10 or any other constant as indicated at step 106, and subtracted from 1.000 as indicated at step 108. As indicated at step 110, the ~~The~~ values for seven and eight are simply assigned as .999 and 1.000, respectively, so that a value greater than .990 (the value for n=6) is produced. However, this is a heuristic construct, as there is insufficient scale for representing how much better a match with seven or eight bits (as compared to a match with six bits) really is.

Please replace the paragraph beginning at page 25, line 12, which starts with "It is noted that the foregoing" with the following amended paragraph:

It is noted that the foregoing has outlined some of the more pertinent objects and embodiments of the present invention. The concepts of this invention may be used for many applications, as discussed above. Thus, although the description is made for particular arrangements and methods ~~as an examples~~ for exemplary purposes, the intent and concept of the invention is suitable and applicable to other arrangements and applications. It will be clear to those skilled in the art that other modifications to the disclosed embodiments can be effected without departing from the spirit and scope of the invention. The described embodiments ought to be construed to be merely illustrative of some of the more prominent features and applications of the invention. Other beneficial results can be realized by applying the disclosed invention in a different manner or modifying the invention in ways known to those familiar with the art. Thus, it should be understood that the embodiments ~~has~~ have been provided as an example and not as a limitation. The scope of the invention is defined by the appended claims.